BORANG PENGESAHAN STATUS TESIS

JUDUL: INVESTIGATING GOLDDREAM ANDROID MALWARE BEHAVIOR THROUGH DYNAMIC ANALYSIS

SESI PENGAJIAN: SESI 2012/2013

Saya LOW JUN KEAT mengaku membenarkan tesis Projek Sarjana Muda ini disimpan di Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dengan syarat-syarat kegunaan seperti berikut:

1. Tesis dan projek adalah hak milik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. ** Sila tandakan (/)

_____________ SULIT (Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)

_____________ TERHAD (Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

/ TIDAK TERHAD

(TANDATANGAN PENULIS) (TANDATANGAN PENYELIA)

Alamat tetap : 924, Permatang Bendahari 13110 Penaga S.P.U Pulau Pinang, Malaysia

DR. SITI RAHAYU SELAMAT Nama Penyelia

Tarikh: 30/8/2013 Tarikh: 30/8/2013

© Universiti Teknikal Malaysia Melaka
INVESTIGATING GOLDDREAM ANDROID MALWARE BEHAVIOR THROUGH DYNAMIC ANALYSIS

LOW JUN KEAT

This report is submitted in partial fulfillment of the requirement for the Bachelor of Computer Science (Computer Networking)

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY
UNIVERSITI TEKNIKAL MALAYSIA MELAKA
2013
DECLARATION

I hereby declare this project report entitled

INVESTIGATING GOLDDREAM ANDROID MALWARE BEHAVIOR
THROUGH DYNAMIC ANALYSIS

is written by me and is my own effort and that no part has been plagiarized without citations.

STUDENT : ___________________________ Date: 30/8/2013

(LOW JUN KEAT)

SUPERVISOR: ________________________ Date: 30/8/2013

( DR. SITI RAHAYU SELAMAT )
DEDICATION

This work is dedicated to my beloved family and siblings, who passed on a love of reading and respect for education.

To my supportive friends, my supervisor and all lecturers, thank you so much for assist and help.
ACKNOWLEDGEMENTS

Thanks the god for giving me the opportunity to complete this Final Year Project which is titled Investigating GoldDream Android Malware Behavior Through Dynamic Analysis.

Firstly, I would like to thanks to, my supervisor, DR. SITI RAHAYU SELAMAT for her technical guidance and valuable time in assisting us in the accomplishing my project. I also want to thanks to all lecturers for their cooperation during completing my final year project by giving valuable information, suggestions and guidance in the compilation and preparation of this report.

Lastly, deepest thanks and appreciation to my parents, family and friends for their understanding, cooperation and full of support for the report completion, from the beginning till the end of this project. Also thanks to all of my friends and everyone, that has been contributed by supporting my work and helps myself during the final year project progress until it is fully completed.
ABSTRACT

In recent year, the growing of Android user are become popular. Unfortunately, as Android is getting more popular, at the same time, it cause the growing of the mobile malware. As the malware are growth rapidly, the current problem is difficulty on detecting and identifying the behavior of android malware. Thus, the aim of this project is to investigate GoldDream Android malware behavior through dynamic analysis. This project used software and hardware tool such as Wireshark for capturing network traffic, emulator for running the android malware applications and Windows 7 operating system as a platform in order to complete the analysis. Hence, the emulator, Wireshark and other tools are installed in Windows 7 operating system which the experiment is executed and data is collected. The objective of this project is to investigate the parameter of android malware's behavior, generate the attack pattern of android malware and formulate the procedure of extracting the attack pattern. The project are start with literature review, analysis, design and implementation, finally is evaluate and testing. In the end of project, the general attack pattern of GoldDream malware is generated based on its basic attack model and its attributes. Then, the GoldDream attack pattern extraction script is developed base on the attack pattern.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>SUBJECT</th>
<th>PAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DECLARATION</td>
<td>i</td>
</tr>
<tr>
<td></td>
<td>DEDICATION</td>
<td>ii</td>
</tr>
<tr>
<td></td>
<td>ACKNOWLEDGMENT</td>
<td>iii</td>
</tr>
<tr>
<td></td>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td></td>
<td>ABSTRAK</td>
<td>v</td>
</tr>
<tr>
<td></td>
<td>TABLE OF CONTENTS</td>
<td>vi</td>
</tr>
<tr>
<td></td>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
</tbody>
</table>

1 INTRODUCTION
1.1 Project Background 1
1.2 Problem Statements 2
1.3 Project Questions 2
1.4 Project Objectives 3
1.5 Project Contribution 4
1.6 Scope 4
1.7 Project Significant 5
1.8 Report Organization 5
1.9 Summary 6

2 LITERATURE REVIEW
2.1 Introduction 7
2.2 Android 8
    2.2.1 Definition 8
2.2.2 Architecture

2.3 Malware

2.3.1 Malware Definition
2.3.2 Malware Distribution
2.3.3 Malware Classification
2.3.4 Basic Malware Attack Model
2.3.5 Trojan Horse Definition
2.3.6 Trojan Horse Classification
2.3.7 Android Malware

2.4 Analysis Technique

2.4.1 Static Analysis
2.4.2 Dynamic Analysis

2.5 Parameter

2.5.1 Network Traffic
2.5.2 System Call

2.6 Attack Pattern

2.7 Summary

3 METHODOLOGY

3.1 Introduction

3.2 Project Methodology

3.2.1 Phase I: Literature Review
3.2.2 Phase II: Analysis
3.2.3 Phase III: Design & Implementation
3.2.4 Phase V: Testing & Evaluation

3.3 Project Schedule & Milestones

3.3.1 Gantt Chart
3.3.2 Milestones

3.4 Summary

4 DESIGN AND IMPLEMENTATION

4.1 Introduction
4.2 GoldDream Analysis Approach 30
   4.2.1 Setup Network 31
   4.2.2 Setup Environment 33
   4.2.3 Malware Activation & Data Collection 34
   4.2.4 Data Analysis 37
4.3 GoldDream attack pattern extraction 55
   Scripting Design
   4.3.1 General Design for GoldDream attack pattern extraction Script 55
   4.3.2 GoldDream attack pattern extraction Process Script 56
4.4 Summary 60

5 TESTING AND RESULT ANALYSIS 61
5.1 Introduction
5.2 Test Planning 61
   5.2.1 Test Organization 61
   5.2.2 Test Environment 62
   5.2.3 Test Schedule 65
5.3 Test Strategy 65
   5.3.1 Unit Testing 66
   5.3.2 Integrate Testing 66
5.4 Test Design 66
   5.4.1 Test Data Set 66
5.5 Test Result 67
   5.5.1 Test Result for System Call 67
   5.5.2 Test Result for Network Traffic 71
5.6 Result Analysis 73
   5.6.1 System Call Result Analysis 73
   5.6.2 Network traffic Result Analysis 78
5.7 Discussion of Result Analysis 82
5.7.1 Capturing Phone Call Log and Sms Malware Activity
5.7.2 Sending Information to Server
5.7.3 General GoldDream Attack Pattern

5.8 Summary

6 CONCLUSION
6.1 Introduction
6.2 Project Summarization
6.3 Project Contribution
6.4 Project Limitation
6.5 Further Project

REFERENCES

APPENDIX
5.5 Datasets summarization

5.6 Comparison of system call result between script output and manual analysis for DS1

5.7 Comparison of system call result between script output and manual analysis for DS2

5.8 Comparison of system call result between script output and manual analysis for DS3

5.9 Comparison of system call result between different Dataset

5.10 Comparison of network traffic result between script output and manual analysis for DS1

5.11 Comparison of network traffic result between script output and manual analysis for DS2

5.12 Comparison of network traffic result between script output and manual analysis for DS3

5.13 Comparison of network traffic result between different Dataset
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Summary of problem statement</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Summary of project questions</td>
<td>2</td>
</tr>
<tr>
<td>1.3</td>
<td>Summary of project objectives</td>
<td>3</td>
</tr>
<tr>
<td>1.4</td>
<td>Summary of project contributions</td>
<td>4</td>
</tr>
<tr>
<td>2.1</td>
<td>Malware classification and Definition</td>
<td>11</td>
</tr>
<tr>
<td>2.2</td>
<td>Trojan Horse classification and Definition</td>
<td>13</td>
</tr>
<tr>
<td>2.3</td>
<td>Android malware type and Definition</td>
<td>15</td>
</tr>
<tr>
<td>2.4</td>
<td>Description of each field in IP Header</td>
<td>20</td>
</tr>
<tr>
<td>2.5</td>
<td>Description of each field in TCP Header</td>
<td>21</td>
</tr>
<tr>
<td>3.1</td>
<td>PSM 1 Milestones</td>
<td>29</td>
</tr>
<tr>
<td>4.1</td>
<td>Collected user and devices information</td>
<td>48</td>
</tr>
<tr>
<td>4.2</td>
<td>Network traffic attributes</td>
<td>50</td>
</tr>
<tr>
<td>4.3</td>
<td>System call attributes</td>
<td>51</td>
</tr>
<tr>
<td>5.1</td>
<td>Test organization</td>
<td>62</td>
</tr>
<tr>
<td>5.2</td>
<td>Hardware Requirements</td>
<td>63</td>
</tr>
<tr>
<td>5.3</td>
<td>Software Requirements</td>
<td>63</td>
</tr>
<tr>
<td>5.4</td>
<td>Test Schedule</td>
<td>65</td>
</tr>
<tr>
<td>FIGURE</td>
<td>TITLE</td>
<td>PAGE</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2.1</td>
<td>Operational framework: Literature review phase</td>
<td>8</td>
</tr>
<tr>
<td>2.2</td>
<td>Android Architecture</td>
<td>9</td>
</tr>
<tr>
<td>2.3</td>
<td>Malware distribution in 2013</td>
<td>10</td>
</tr>
<tr>
<td>2.4</td>
<td>Malware classification</td>
<td>10</td>
</tr>
<tr>
<td>2.5</td>
<td>General Malware attack pattern</td>
<td>12</td>
</tr>
<tr>
<td>2.6</td>
<td>Basic Malware attack model</td>
<td>12</td>
</tr>
<tr>
<td>2.7</td>
<td>Trojan Horse classification</td>
<td>13</td>
</tr>
<tr>
<td>2.8</td>
<td>Percentage of mobile operating system infected by malware</td>
<td>14</td>
</tr>
<tr>
<td>2.9</td>
<td>Analysis technique</td>
<td>17</td>
</tr>
<tr>
<td>2.10</td>
<td>IP header</td>
<td>19</td>
</tr>
<tr>
<td>2.11</td>
<td>TCP header</td>
<td>19</td>
</tr>
<tr>
<td>3.1</td>
<td>Project Methodology</td>
<td>25</td>
</tr>
<tr>
<td>3.2</td>
<td>Project Gantt Chart</td>
<td>28</td>
</tr>
<tr>
<td>4.1</td>
<td>Analysis Approach</td>
<td>30</td>
</tr>
<tr>
<td>4.2</td>
<td>Physical design of malware analysis environment</td>
<td>31</td>
</tr>
<tr>
<td>4.3</td>
<td>Logical design of malware analysis environment</td>
<td>32</td>
</tr>
<tr>
<td>4.4</td>
<td>Process of collect network traffic</td>
<td>34</td>
</tr>
</tbody>
</table>
4.5 Process of collect system call
4.6 Data analysis Approach
4.7 Collected network traffic
4.8 Content of network traffic
4.9 Capture of phone call log
4.10 Capture of received message content
4.11 Location of zjphonecall.txt and zjsms.txt
4.12 Content of zjphonecall.txt
4.13 Content of zjsms.txt
4.14 Process of sending user and devices information
4.15 Collect user and devices information
4.16 Resolve domain name to IP address
4.17 Establish connection to server
4.18 Process flow of capture the phone log and sms
4.19 Basic Attack Model
4.20 GoldDream attack Pattern
4.21 General design for goldDream attack pattern extraction script
4.22 Overall process of GoldDream attack pattern extraction script
4.23 Phone call and sms identification
4.24 Information stealing identification
4.25 Extraction of malicious network traffic payload
5.1 Test environment design
5.2 Result of capture phone call log and sms from DS 1 68
5.3 Result of capture phone call log and sms from DS 2 68
5.4 Result of capture phone call log and sms from DS 3 69
5.5 Result of send user and devices information to server from DS 1 70
5.6 Result of send user and devices information to server from DS 2 70
5.7 Result of send user and devices information to server from DS 3 71
5.8 Result of network traffic from DS 1 72
5.9 Result of network traffic from DS 2 73
5.10 Result of network traffic from DS 3 73
5.11 Example captured data 83
5.12 General format of captured data 83
5.13 Example location of store text file 84
5.14 General pattern of file location 84
5.15 Method to save captured data 84
5.16 Request url 86
5.17 General url pattern 86
5.18 Proposed General goldDream Attack Pattern 87
CHAPTER 1

INTRODUCTION

1.1 Project Background

In recent years, the growing of smartphone is becoming popular. There are various operating systems in smartphones such as iOS, Window, Android as well as Symbian. Android is the most popular (40.9%) compared to Apple iOS (40.5%), Blackberry (8.9%) (Arlotta, 2013). Unfortunately, as Android is getting more popular, at the same time, it causes the growing of mobile malware.

Malware is also known as malicious software used or created by attackers to damage or destroy computer operation, steal sensitive information and break into private computer systems. It can emerge as a collection of code, script, active content and embedded in other applications (McMahon, 2013).

The malware grows rapidly, it is a need to take effectively defend against the malware. In order to create a method to against the malware, we need to study, analyse and investigate the malware. Malware analysis is an activity in which carry out by reverse engineering the malware and investigate on code structure, operation and functionality (Varghese, 2011).

The goal of this project is to understand the behavior of an Android malware. In addition, Android OS is a popular environment for mobile malware. It needs to take an action to overcome it before it gets serious. However, we need to understand how it works before we can defend it.
Therefore, this project will use dynamic analysis to analyse the malware. It will focuses on the behavior of attack, determine how and what it gets installed, how it run, what background process had been create, which port are using to communicate, who they are communicate to, etc. (Distler, 2007). In this project, the parameter such as system call and network traffic will be investigated.

1.2 Problem Statement

Malware will spread widely, rapidly, will embed in other software and some may encrypted the network traffic. This characteristic causes the difficulty to detect and identify the malware (Yen, 2011). The Project Problem (PP) is summarized into Table 1.1.

Table 1.1 Summary of problem statement

<table>
<thead>
<tr>
<th>No</th>
<th>Research Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP1</td>
<td>Difficulty on detecting and identifying the behavior of Android malware</td>
</tr>
</tbody>
</table>

1.3 Project Questions

Three Project Questions (PQ) is constructed to identify the problem statement as discussed in previous section is depicted in Table 1.2.

Table 1.2 Summary of project questions

<table>
<thead>
<tr>
<th>RP</th>
<th>RQ</th>
<th>Research Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP1</td>
<td>PQ1</td>
<td>What is the parameter use to study the behaviour of Android malware?</td>
</tr>
<tr>
<td></td>
<td>PQ2</td>
<td>What is the behavior of android malware?</td>
</tr>
<tr>
<td></td>
<td>PQ3</td>
<td>What is the procedure of extract the behavior?</td>
</tr>
</tbody>
</table>
PQ1: What is the parameter used to study the behavior of Android malware?

This project question is to analyse which parameter is suitable to use to study the behavior of Android malware. Because different types of malware may infect different parameters, it is important to analyse which parameter should be used.

PQ2: What is the behavior of Android malware?

This project question is to study and identify which technique is suitable to use to collect the data that are used to identify the behavior.

PQ3: What is the procedure of extracting the behavior?

This project question is to find out how to extract the behavior and generate the attack pattern automatically.

1.4 Project Objectives

Based on the project questions formulated in the previous section, appropriate project objectives (PO) are developed as follows:

<table>
<thead>
<tr>
<th>RP</th>
<th>RQ</th>
<th>RO</th>
<th>Research Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP1</td>
<td>PQ1</td>
<td>PO1</td>
<td>To investigate the parameter of Android malware's behavior</td>
</tr>
<tr>
<td></td>
<td>PQ2</td>
<td>PO2</td>
<td>To generate the attack pattern of Android malware</td>
</tr>
<tr>
<td></td>
<td>PQ3</td>
<td>PO3</td>
<td>To formulate the procedure of extracting the attack pattern</td>
</tr>
</tbody>
</table>

PO 1: To investigate the parameter of Android malware's behavior

In order to analyse the Android malware, first we must identify what parameter will be used to analyse the malware. Different types of malware may have different types of parameter to inspect.
PO 2: To generate the attack pattern of Android malware

After determine the parameter use to analyses the malware, the next step is to collect data and analyses the data to identify the behavior in order to generate attack pattern.

PO3: To formulate the procedure of extracting the attack pattern

After generate the attack pattern of malware, then will formulate the procedure and develop a script to extract the attack pattern automatically from raw data.

1.5 Project Contributions

The contribution of this project are summarized in Table 1.4

<table>
<thead>
<tr>
<th>RP</th>
<th>RQ</th>
<th>RO</th>
<th>RC</th>
<th>Project Contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP1</td>
<td>PQ1</td>
<td>PO1</td>
<td>PC1</td>
<td>The parameter use to analyses android malware's behavior</td>
</tr>
<tr>
<td></td>
<td>PQ2</td>
<td>PO2</td>
<td>PC2</td>
<td>The attack pattern of android malware</td>
</tr>
<tr>
<td></td>
<td>PQ3</td>
<td>PO3</td>
<td>PC3</td>
<td>The script to extract android attack pattern</td>
</tr>
</tbody>
</table>

1.6 Project Scope

The project will be focussed on:

a. GoldDream malware
b. System call and network traffic parameter
c. Using dynamic analysis
d. Develop script to extract the attack pattern
1.7 Project Significant

The attack pattern of GoldDream will help developer in develop a method or software to protect the system from GoldDream malware.

1.8 Report organization

This report consist of six chapter namely Chapter 1: Introduction, Chapter 2: Literature Review, Chapter 3: Methodology, Chapter 4: Design and Implementation, Chapter 5: Testing and Result Analysis and Chapter 6: Conclusion.

Chapter 1: Introduction
This chapter will discuss about introduction, project background, research problem, research question, research objective, scope, project significant and report organization.

Chapter 2: Literature Review
This chapter will explain related work of this project, such as Android, malware, analysis technique and parameter.

Chapter 3: Methodology
This chapter will explain the method use to analyse the malware and organise the sequence of project work in phase by phase.

Chapter 4: Design and Implementation
This chapter will introduce the software and hardware use in this project, environment setup, implementation of malware as well as the data collected.

Chapter 5: Testing and Result Analysis
This chapter will analyse the collected data and carry out the scripting proposed to support the evidence.
Chapter 6: Conclusion
This chapter will concludes and discussed the finding, limitations, contribution and the future work of the project.

1.9 Summary

In this chapter, problem statement, questions and objective of the projects are clearly identified. The next chapter, Chapter 2 will discuss the related work of this project.
CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

In chapter 1, the problem statement, questions and objective of the projects are clearly identified. For this chapter, the main topics are literature review will be discussed. The aim of this chapter is to review several issues related with this project, such as Android, malware, analysis technique and parameter as depicted in Figure 2.1.

![Operational framework: Literature review phase](image-url)
Figure 2.1 shows the topic will be discussed in the following sections. Further information on Android, malware, analysis technique and system parameter issues are gathered. During the Literature Review phase, the relevant literature in journals, articles, thesis, technical reports, books, websites and other academic sources are reviewed. The 5 main issues to discuss in literature review is Android, malware, analyses technique, system parameter and attack pattern.

2.2 Android

In this section, the definition and the architecture of android is discussed.

2.2.1 Definition

Android is a Linux-based operating system (Katsarakis, 2012) that mainly designed for touch screen mobile device such as tablet and smart phone. Android, Inc. was the first who start to develop Android and it is financially backed by Google, but in 2005 it is bough by Google (Elgin and Ben, 2005). Android is open source and Google release the code under the Apache License in september 2008 (Katsarakis, 2012).

2.2.2 Architecture

The Android architecture structural diagram is shown in Figure 2.2. The Android architecture are consist of 5 layer, which is the lowest layer Linux kernel layer, native libraries, the Android Runtime, the application framework layer and application layer is on the top layer (Brahler, 2010).